

ABSTRACT OF THE DISCLOSURE

Speech recognition performance is improved without changing a speech recognition engine. A speech data generation section generates, from speech data for which speech recognition is to be performed, a plurality of pieces of speech data whose starting positions of the non-speech regions differ. A speech recognition engine performs speech recognition by using each of the pieces of speech data. A totaling/comparison section provides the most numerous recognized result from among a plurality of obtained recognized results.